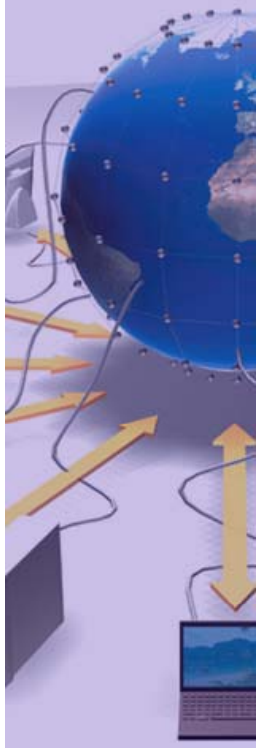
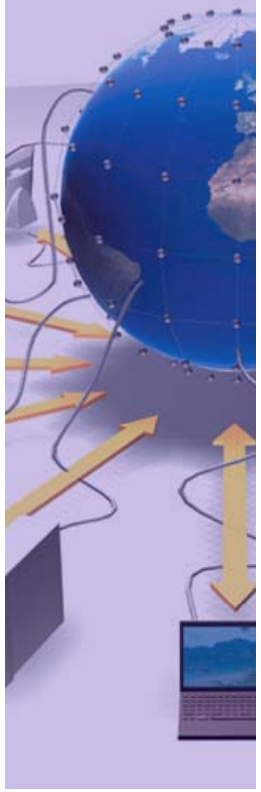


# CCRC08 ATLAS Post Mortem

Simone Campana
IT/GS/EIS

CERN IT Department
CH-1211 Genève 23
Switzerland
**www.cern.ch/it**

# CCRC08 phase I and II

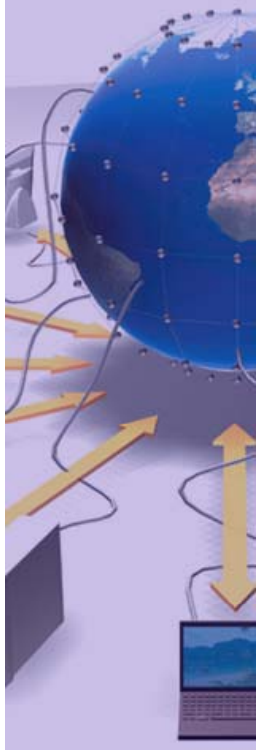


- **The aim of CCRC08 is to test al experiments activities together**

- **CCRC08 Phase I:**
  - Mostly a test of SRMv2 installation/configuration
    - (functionality)
  - For ATLAS, very short exercise
    - Concurrent with FDR in week I and II
- **CCRC08 Phase II:**
  - Tests carried along for the all month
    - No overlap with FDR (1st week of June)
    - **CCRC08 ONLY during week days**
    - Cosmic data during the weekend (commissioning and M7)
  - Focused on data distribution
    - T0->T1, T1->T1, T1->T2
  - Very demanding metrics
    - More than you will need to do during 2008 data taking

CERN IT Department
CH-1211 Genève 23
Switzerland
**www.cern.ch/it**

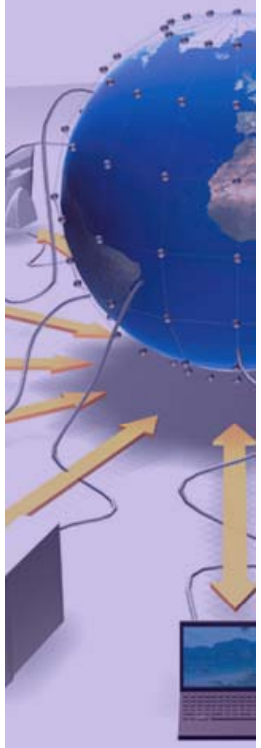
# Recurring concepts



- **"The load generator"**
  - Agent Running at the T0, generates "fake data" as if they were coming from the detector.
    - Fake reconstruction jobs **run** in LSF
    - Dummy files (not compressible) **stored** on CASTOR
    - Files organized in datasets and **registered** in LFC, dataset **registered** in ATLAS DDM Central Catalog
    - Generally big files (from 300MB to 3GB)

- **The "Double Registration" problem**
  - The file is transferred correctly to site X and registered in LFC
  - "Something" goes wrong and the file is replicated again
    - Another entry in LFC, same GUID, different SURL

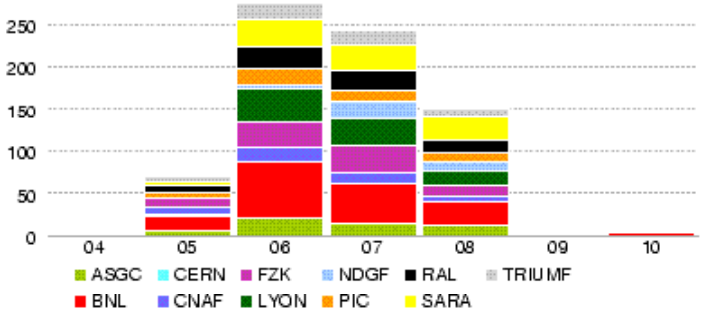
# Week-1: DDM Functional Test



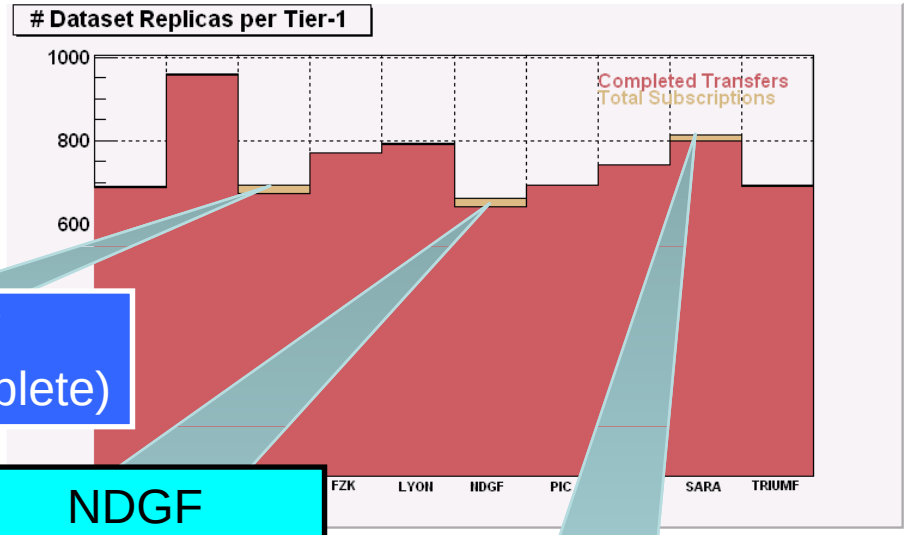
- Running Load Generator for 3 days at 40% of nominal rate
- Dataset subscribed to T1 DISK and TAPE endpoints
  - RAW data subscribed according to ATLAS MoU shares (TAPE)
  - ESD subscribed ONLY at the site hosting the parent RAW datasets (DISK)
    - In preparation for T1-T1 test of Week 2
  - AOD subscribed to every site (DISK)
- No activity for T2s in week 1
- **Metrics**:
  - Sites should hold a complete replica of 90% of subscribed datasets
  - Dataset replicas should be completed at sites within 48h from subscription

CERN IT Department
CH-1211 Genève 23
Switzerland
**www.cern.ch/it**

# Week-1 Results




*CCRC08. May 2008 (week 19)*




CNAF
(97% complete)

NDGF
(94% complete)

SARA
97% complete

**Temporary failure (disk server) treated as permanent by DDM. Transfer not retried**

**Problematic throughput to TAPE**

**Limited resources:**
**1 disk buffer in front of 1 tape drive.**
**Only 4 active transfers allowed.**
**Clashes with FTS configuration (20 transfers).**

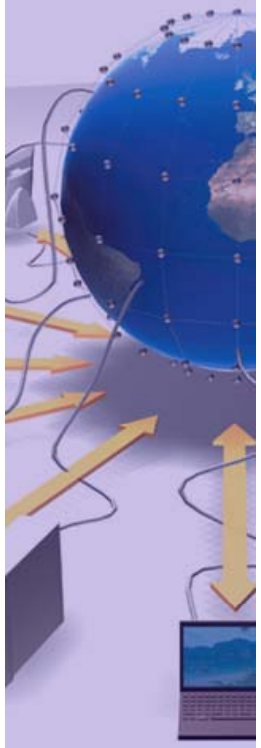
**Competition with other ("production") transfers**

**"Double Registration" problem**

**Slow transfer time out (changed from 600 to 3000s) Storage fails to cleanup the disk pool after the entry of the file was removed from the namespace … disk full**

CERN IT Department
CH-1211 Genève 23
Switzerland
**www.cern.ch/it**

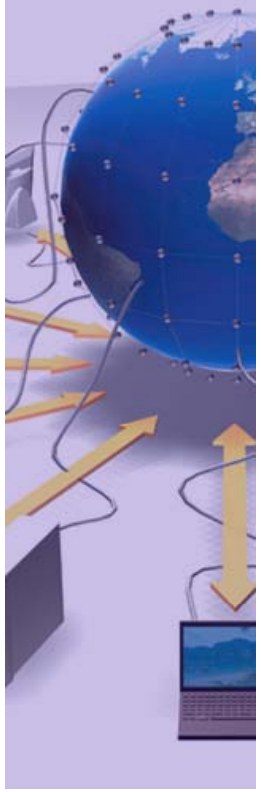
# Week-2: T1-T1 test



- Replicate ESD of week 1 from “hosting T1” to all other T1s.
  - Test of the full T1-T1 transfer matrix
  - FTS at destination site schedules the transfer
  - Source site is always specified/imposed
    - No chaotic T1-T1 replication … not in the ATLAS model.
- Concurrent T1-T1 exercise from CMS
  - Agreed in advance

CERN IT Department
CH-1211 Genève 23
Switzerland
**www.cern.ch/it**

# Week-2: T1-T1 test

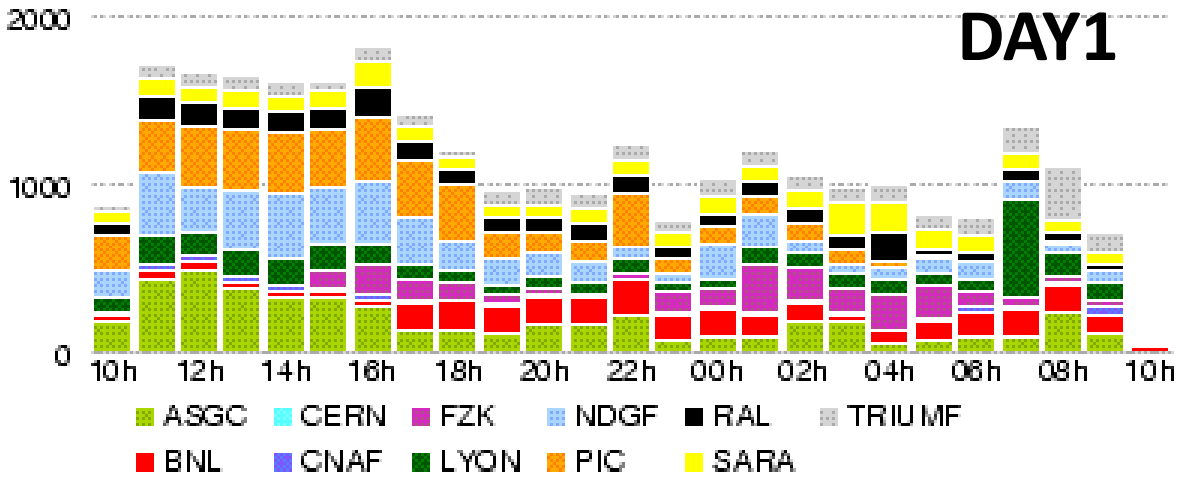


- Dataset sample to be replicated:
  - 629 datasets corresponding to 18TB of data
    - For NL, SARA used as source, NIKHEF as destination

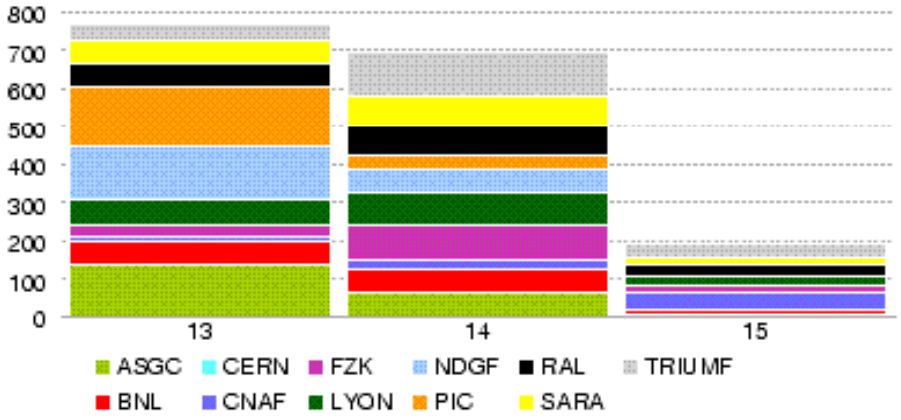
- **Timing and Metrics:**
  - Subscriptions to every T1 at 10 AM on May 13th
    - All in one go … will the system throttle or collapse?
  - Exercise finishes at 2 PM on May 15th
  - For every “channel” (T1-T1 pair) 90% of datasets should be completely transferred in the given period of time.
    - Very challenging: 90MB/s import rate per each T1!

CERN IT Department
CH-1211 Genève 23
Switzerland
**www.cern.ch/it**

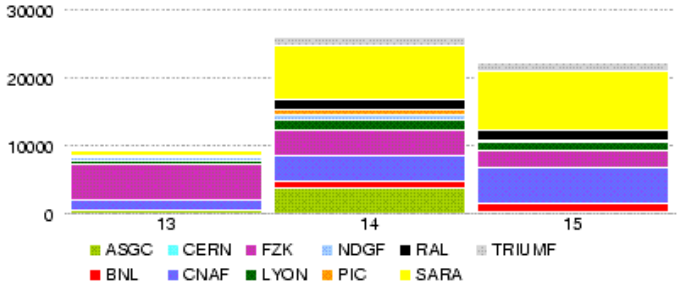
# Week-2: Results




## All days (throughput)




## All days (errors)

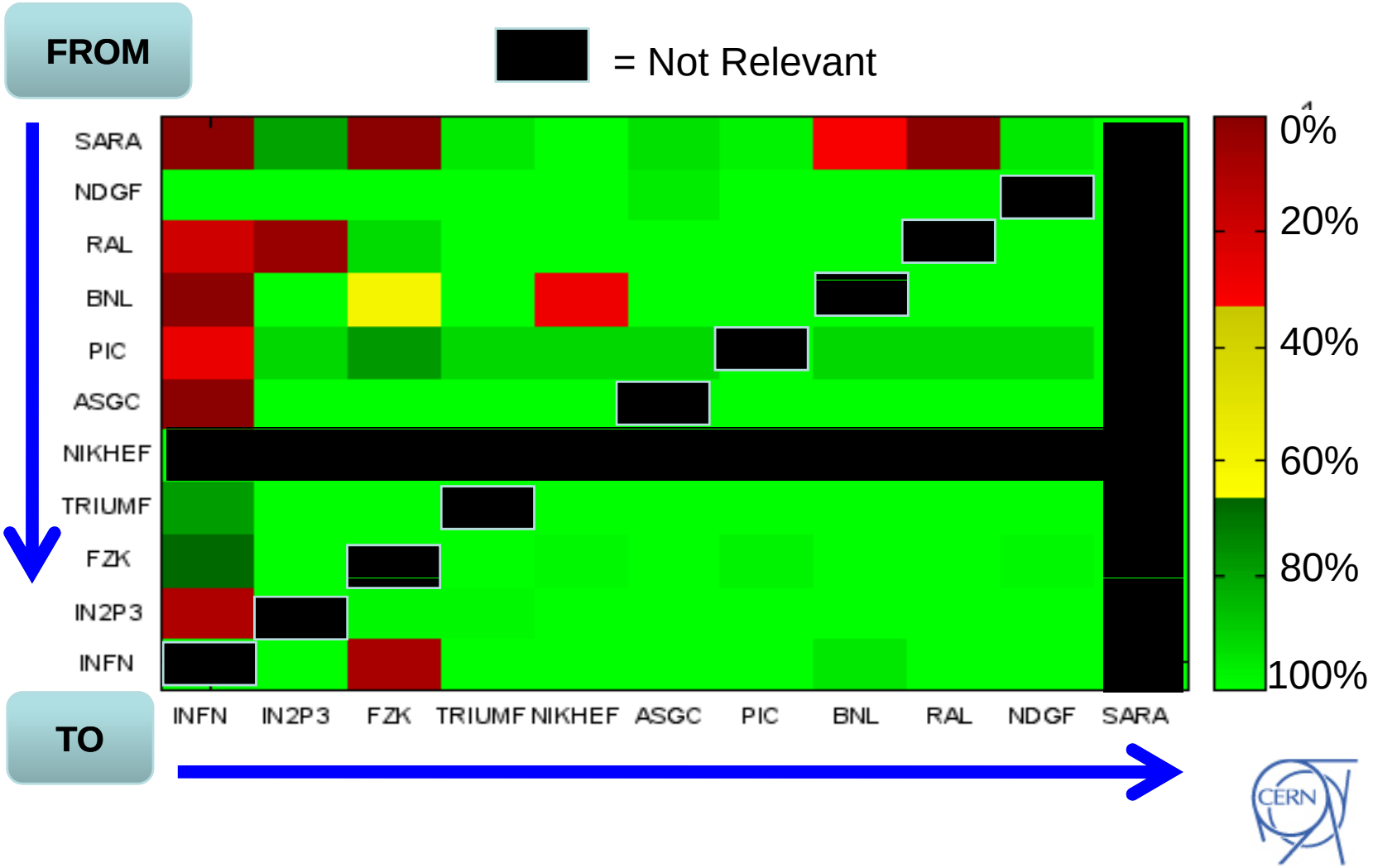



CERN IT Department
CH-1211 Genève 23
Switzerland
**www.cern.ch/it**

# Week-2: Results

CERN IT Department

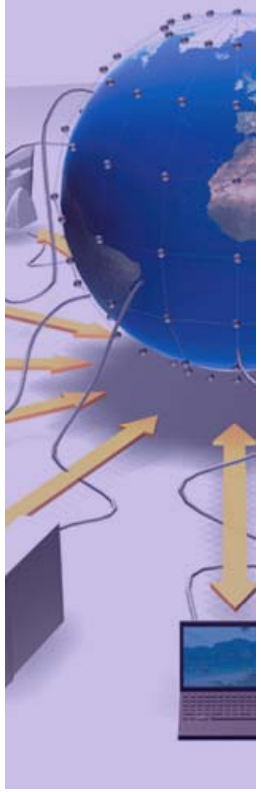
## Fraction of completed dataset




CERN IT Department
CH-1211 Genève 23
Switzerland
**www.cern.ch/it**

# Week-2: site observations in a nutshell



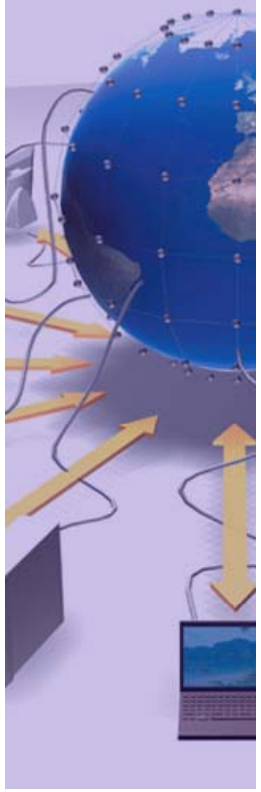


## ***Very highly performing sites***

- **ASGC**
  - 380 MB/s sustained 4 hours, 98% efficiency
  - CASTOR/SRM problems in day 2 dropped efficiency
- **PIC**
  - Bulk of data (16TB) imported in 24h, 99% efficiency
    - With peaks of 500MB/s
  - A bit less performing in data export
    - dCache ping manager unstable when overloaded
- **NDGF**
  - NDGF FTS uses gridFTP2
    - Transfers go directly to disk pools

CERN IT Department
CH-1211 Genève 23
Switzerland
**www.cern.ch/it**

# Week-2: site observations in a nutshell

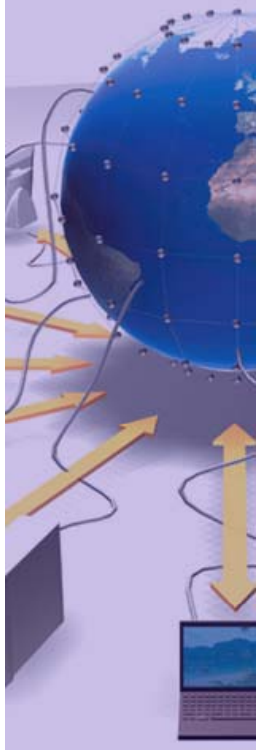


## ***Highly performing sites***

- **BNL**
  - Initial slow ramp-up due to competing production (FDR) transfer
    - Fixed setting FTS priorities
  - Some minor load issues in PNFS
- **RAL**
  - Good dataset completion, slightly low rate
    - Not very aggressive in FTS setting
  - Discovered a RAL-IN2P3 network problem
- **LYON**
  - Some instability in LFC daemon
    - hangs, need restart
- **TRIUMF**
  - Discovered a problem in OPN failover
    - Primary lightpath to CERN failed, secondary was not used.
    - The tertiary route (via BNL) was used instead

CERN IT Department
CH-1211 Genève 23
Switzerland
**www.cern.ch/it**

# Week-2: site observations in a nutshell



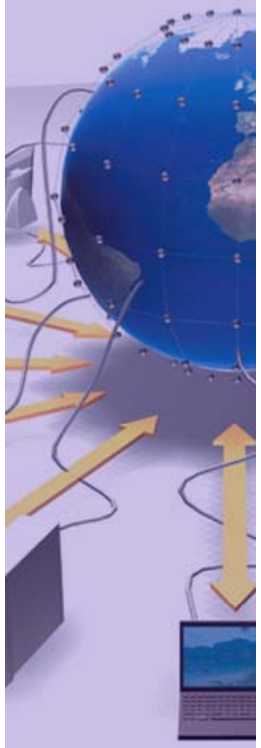
## *“Not very smooth experience” sites*

- **CNAF**
  - problems importing from many sites
  - High load on StoRM gridftp servers
    - A posteriori, understood a peculiar effect in gridFTP-GPFS interaction
- **SARA/NIKHEF**
  - Problems exporting from SARA
    - Many pending gridftp requests waiting on pools supporting less concurrent sessions
    - SARA can support 60 outbound gridFTP transfers
  - Problems importing in NIKHEF
    - DPM pools a bit unbalanced (some have more space)
- **FZK**
  - Overload of PNFS (too many SRM queries). FTS settings..
  - Problem at the FTS oracle backend

CERN IT Department
CH-1211 Genève 23
Switzerland
**www.cern.ch/it**

# FTS files and streams (during week-2)

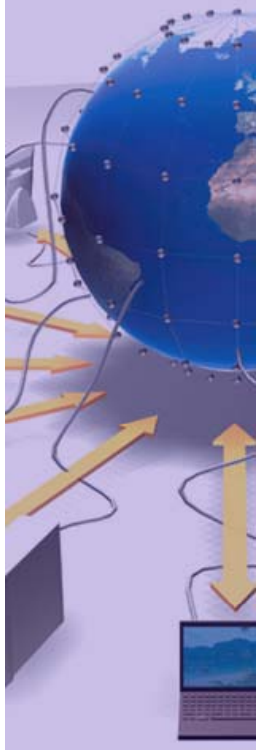
CERN IT Department



| | ASGC | | BNL | | CNAF | | FZK | | Lyon | | NDGF | | NIKHEF | | PIC | | RAL | | Triumf | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | F | S | F | S | F | S | F | S | F | S | F | S | F | S | F | S | F | S | F | S |
| ASGC | - | - | 10 | 10 | 20 | 10 | 10 | 10 | 10 | 10 | 20 | 2 | 10 | 5 | 30 | 10 | 24 | 2 | 10 | 7 |
| BNL | 10 | 20 | - | - | 20 | 10 | 20 | 10 | 10 | 10 | 20 | 2 | 10 | 5 | 30 | 5 | 27 | 1 | 10 | 7 |
| CNAF | 25 | 20 | 10 | 10 | - | - | 20 | 10 | 10 | 10 | 20 | 2 | 10 | 5 | 30 | 5 | 6 | 1 | 10 | 7 |
| FZK | 20 | 10 | 10 | 10 | 20 | 10 | - | - | 10 | 10 | 20 | 2 | 10 | 5 | 30 | 5 | 6 | 1 | 10 | 7 |
| Lyon | 40 | 20 | 10 | 10 | 20 | 10 | 20 | 10 | - | - | 20 | 2 | 10 | 5 | 30 | 5 | 6 | 1 | 10 | 7 |
| NDGF | 10 | 20 | 10 | 10 | 20 | 10 | 20 | 1 | 10 | 10 | - | - | 10 | 5 | 30 | 5 | 6 | 1 | 10 | 7 |
| NIKHEF | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | - | - | 0 | 0 | 0 | 0 | 0 | 0 |
| PIC | 30 | 10 | 10 | 10 | 20 | 10 | 10 | 10 | 10 | 10 | 20 | 2 | 10 | 5 | - | - | 8 | 1 | 10 | 7 |
| RAL | 40 | 20 | 10 | 10 | 20 | 10 | 20 | 10 | 10 | 10 | 20 | 2 | 10 | 5 | 30 | 5 | - | - | 10 | 7 |
| SARA | 40 | 10 | 10 | 10 | 20 | 10 | 10 | 10 | 10 | 10 | 20 | 2 | 10 | 5 | 30 | 5 | 6 | 1 | 20 | 7 |
| Triumf | 15 | 10 | 10 | 10 | 20 | 10 | 20 | 10 | 10 | 10 | 20 | 2 | 20 | 5 | 10 | 5 | 12 | 1 | - | - |

CERN IT Department
CH-1211 Genève 23
Switzerland
**www.cern.ch/it**

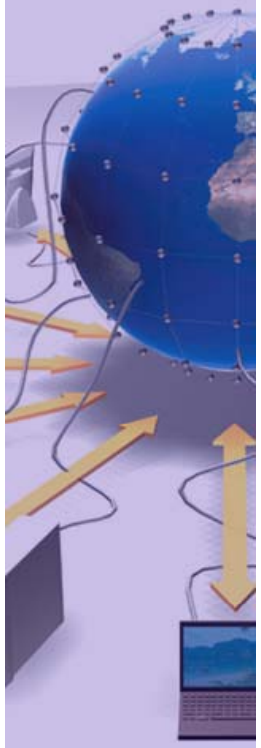
# Week- 2 General Remarks



- Some global tuning of FTS parameters is needed
  - Tune global performance and not single site
  - Very complicated: full matrix must also take into account other VOs.
- FTS at T1s
  - ATLAS would like 0 internal retries in FTS
    - Simplifies Site Services workload, DDM has internal retry anyway (more refined)
    - Could every T1 set this for ATLAS only?
  - Channel <SITE>-NIKHEF has now been set everywhere
    - Or “STAR” channel is deliberately used
  - Would be good to have FTM
    - Monitor transfers in GridView
  - Would be good to have logfiles exposed

CERN IT Department
CH-1211 Genève 23
Switzerland
**www.cern.ch/it**

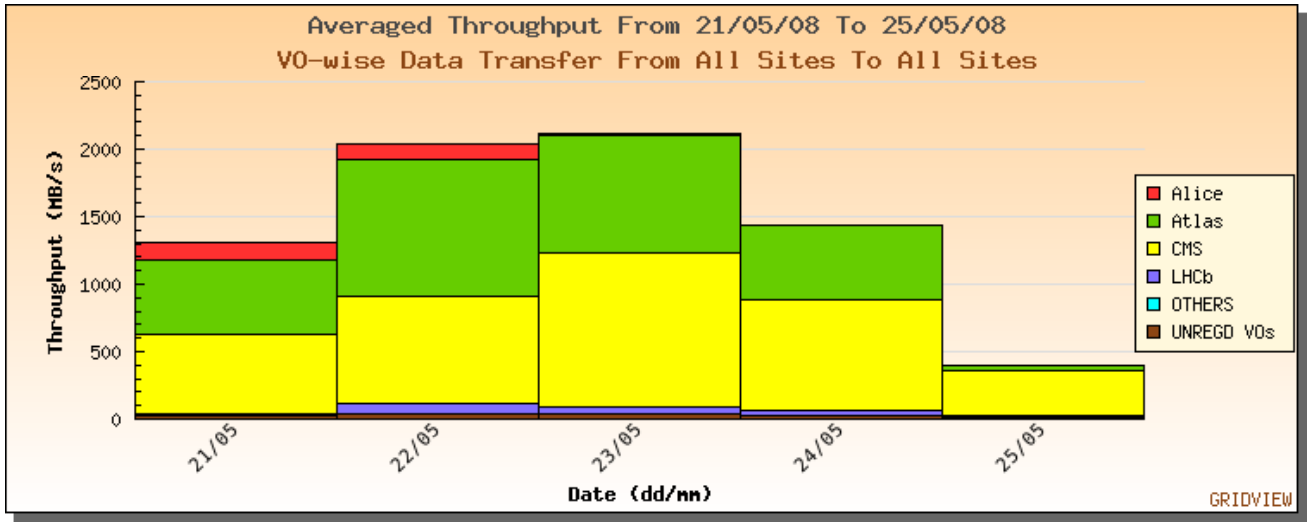
# Week-3: Throughput Test

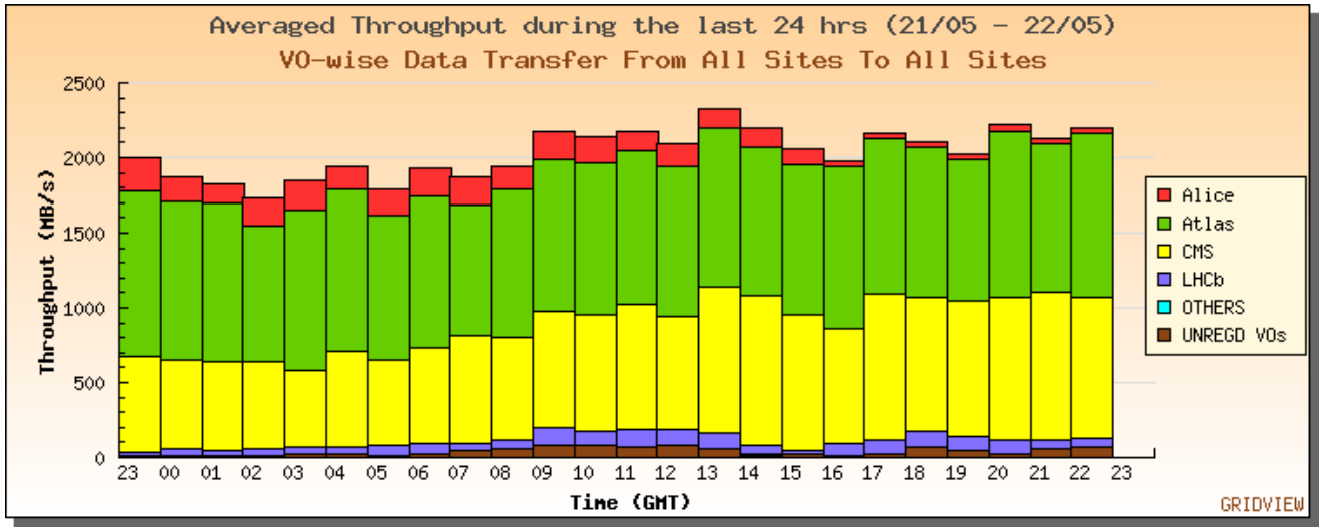


- Simulate data exports from T0 for 24h/day of detector data taking at 200Hz
  - Nominal rate is 14h/day
- No oversubscription
  - Everything distributed according to computing model
    - Weather you get “everything” you are subscribed to or you do not achieve the nominal throughput
- **Timing and Metrics:**
  - Exercise starts on May the 21st at 10AM and ends May the 24th at 10AM
  - Sites should be able to sustain the peak rate for at least 24 hours and the nominal rate for 3 days

CERN IT Department
CH-1211 Genève 23
Switzerland
**www.cern.ch/it**

# Week-3: all experiments

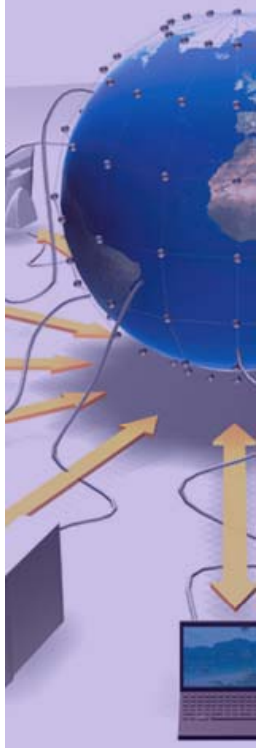



CERN IT Department
CH-1211 Genève 23
Switzerland
**www.cern.ch/it**

# Expected Rates

**RAW: 1.5 MB/event**
**ESD: 0.5 MB/event**
**AOD: 0.1 MB/event**



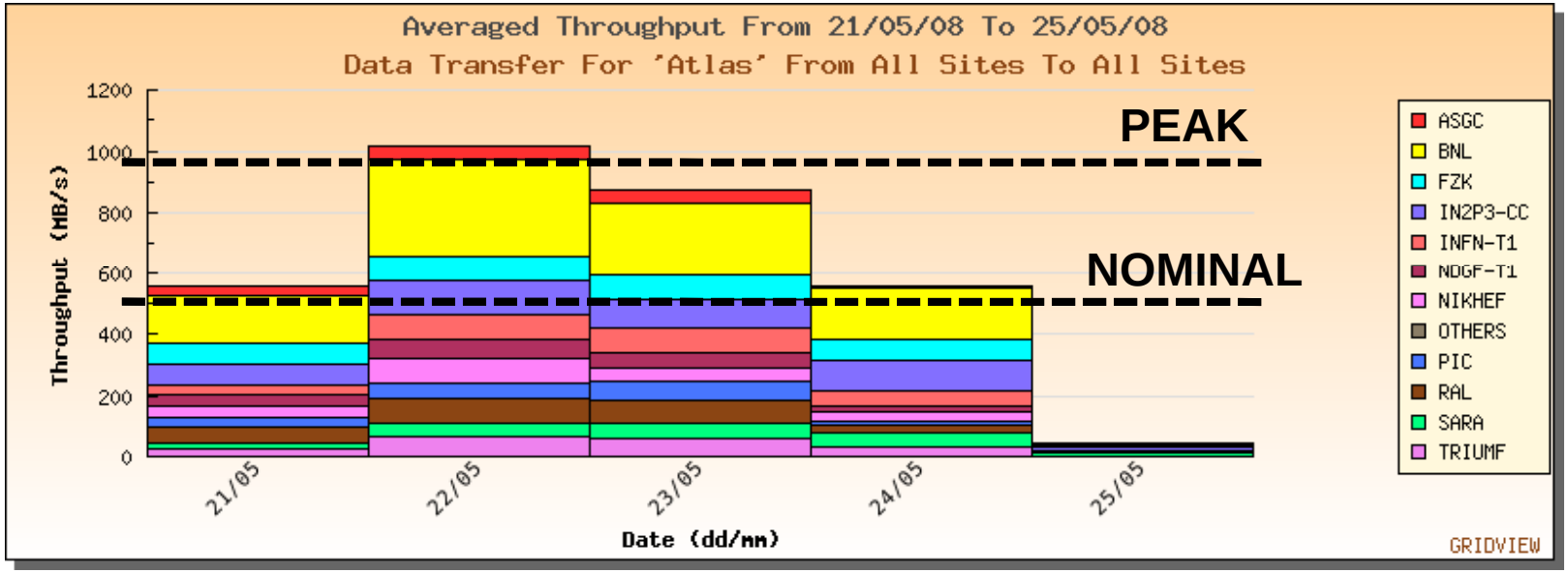
| | | Transfers |
|---|---|---|
| Cloud | Efficiency | Throughput |
| ASGC | 94% | 51 MB/s |
| BNL | 64% | 319 MB/s |
| CERN | 0% | 0 MB/s |
| CNAF | 55% | 77 MB/s |
| FZK | 85% | 118 MB/s |
| LYON | 71% | 120 MB/s |
| NDGF | 63% | 67 MB/s |
| PIC | 75% | 60 MB/s |
| RAL | 84% | 92 MB/s |
| SARA | 43% | 106 MB/s |
| TRIUMF | 79% | 48 MB/s |

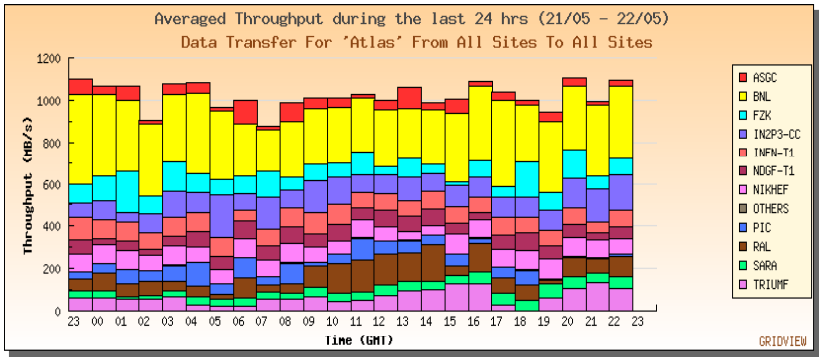
**Snapshot for May 21st**

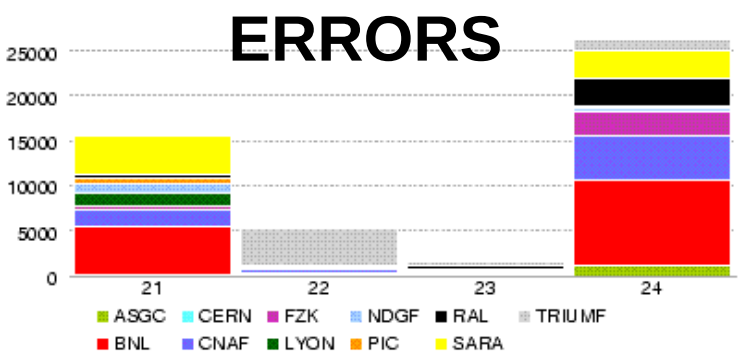
| Rates(MB/s) | TAPE | DISK | Total |
|---|---|---|---|
| BNL | 79.63 | 218.98 | 298.61 |
| IN2P3 | 47.78 | 79.63 | 127.41 |
| SARA | 47.78 | 79.63 | 127.41 |
| RAL | 31.85 | 59.72 | 91.57 |
| FZK | 31.85 | 59.72 | 91.57 |
| CNAF | 15.93 | 39.81 | 55.74 |
| ASGC | 15.93 | 39.81 | 55.74 |
| PIC | 15.93 | 39.81 | 55.74 |
| NDGF | 15.93 | 39.81 | 55.74 |
| Triumf | 15.93 | 39.81 | 55.74 |
| Sum | 318.5 | 696.8 | 1015.3 |

CERN IT Department
CH-1211 Genève 23
Switzerland
**www.cern.ch/it**

# Week-3: Results



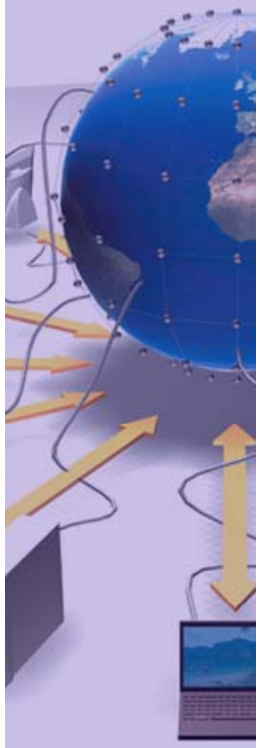




CERN IT Department
CH-1211 Genève 23
Switzerland
**www.cern.ch/it**

# Issues and backlogs

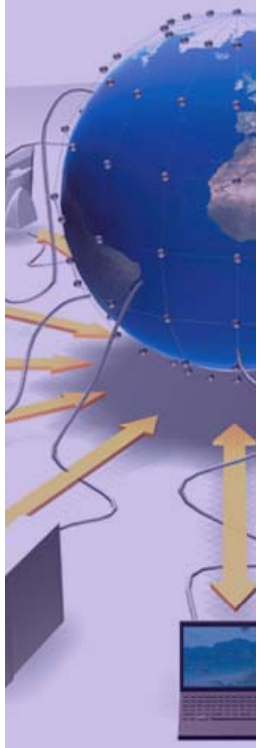
- SRM/CASTOR problems at CERN
  - 21st of May from 16:40 to 17:20 (unavailability)
  - 21st of May from 22:30 to 24:00 (degrade)
  - 24th of May from 9:30 to 11:30 (unavailability)
- Initial problem at RAL
  - UK CA rpm not updated on disk servers
- Initial problem at CNAF
  - Problem at the file system
- Performance problem at BNL
  - backup link supposed to provide 10Gbps was limited at 1Gbps
- 1 write pool at Triumf was offline
  - But dCache kept using it
- SARA TAPE seems very slow but …
  - Concurrently they were writing “production” data
  - In addition they were hit by the double registration problem
  - At the end of the story … they were storing 120 MB/s into tape. Congratulations.



CERN IT Department
CH-1211 Genève 23
Switzerland
**www.cern.ch/it**

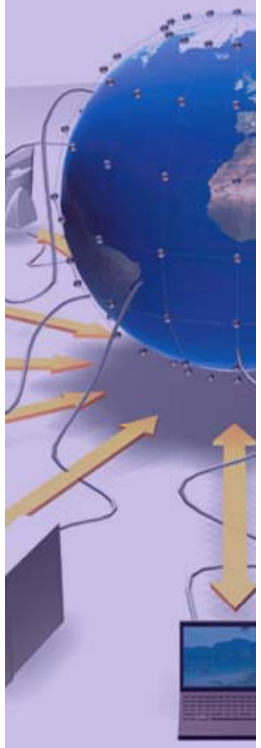
# Week-4: Full Exercise

- The aim is to test the full transfer matrix
  - Emulate the full load T0->T1 + T1->T1 + T1->T2
  - Considering 14h data taking
  - Considering full steam reprocessing at 200Hz
- On top of this, add the burden of Monte Carlo production
  - Attempt to run as many jobs as one can
  - This also means transfers T1->T2 and T2->T1
- Four days exercise divided in two phases
  - First two days: functionality (lower rate)
  - Last two days: throughput (full steam)



CERN IT Department
CH-1211 Genève 23
Switzerland
**www.cern.ch/it**

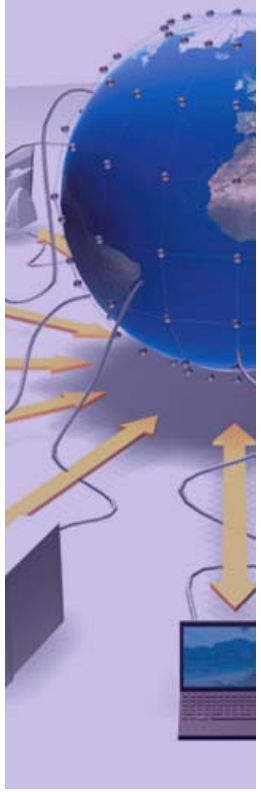
# Week-4: metrics

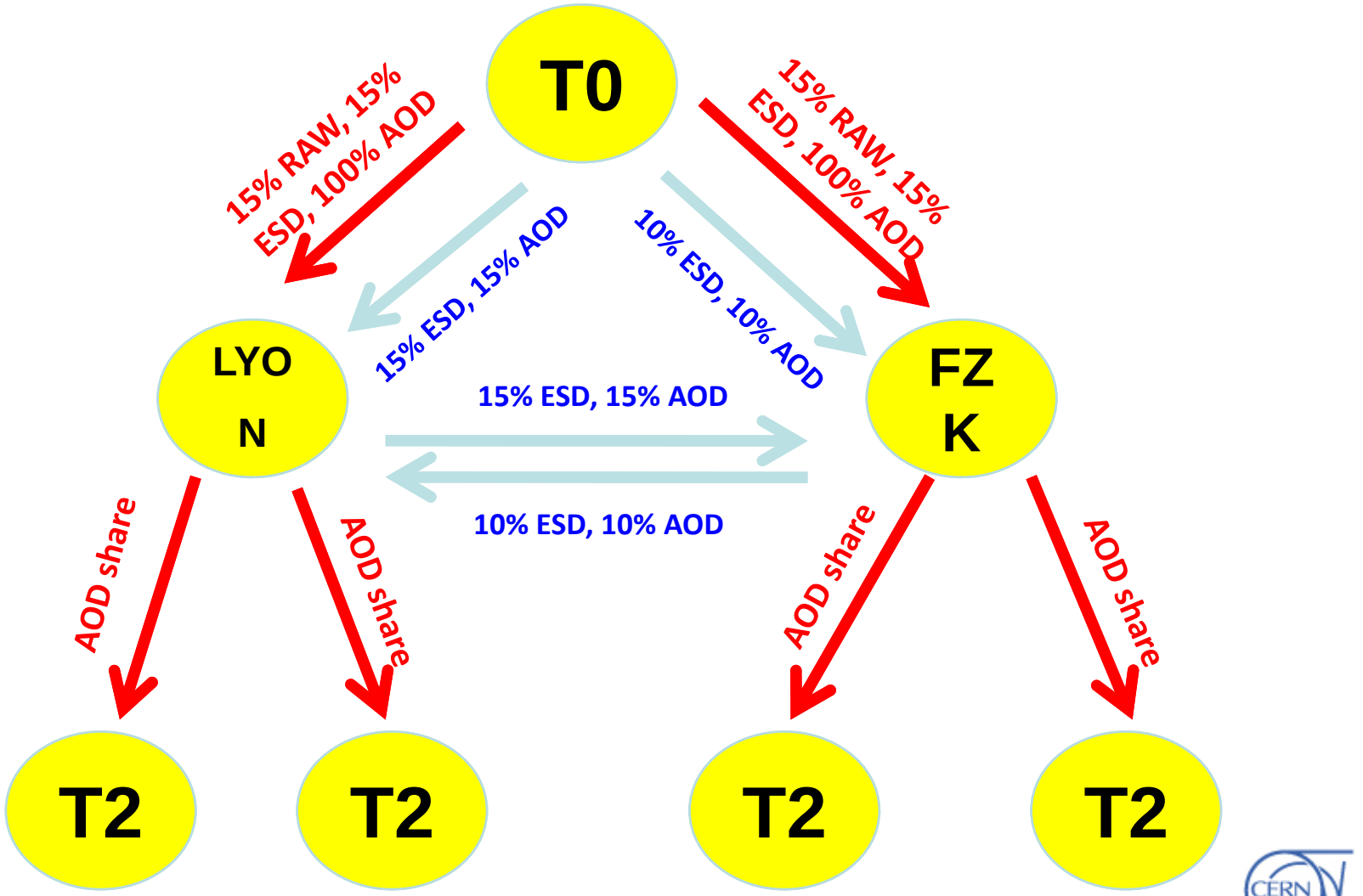


- T0->T1: sites should demonstrate to be capable to import 90% of the subscribed datasets (complete datasets) within 6 hours from the end of the exercise
- T1->T2: a complete copy of the AODs at T1 should be replicated at among the T2s, withing 6 hours from the end of the exercise
- T1-T1 functional challenge, sites should demonstrate to be capable to import 90% of the subscribed datasets (complete datasets) for within 6 hours from the end of the exercise
- T1-T1 throughput challenge, sites should demonstrate to be capable to sustain the rate during nominal rate reprocessing i.e. F*200Hz, where F is the MoU share of the T1. This means:
  - a 5% T1 (CNAF, PIC, NDGF, ASGC, TRIUMF) should get 10MB/s from the partner in ESD and 19MB/s in AOD
  - a 10% T1 (RAL, FZK) should get 20MB/s from the partner in ESD and ~20MB/s in AOD
  - a 15% T1 (LYON, NL) should get 30MB/s from the partner in ESD and ~20MB/s in AOD
  - BNL should get all AODs and ESDs

CERN IT Department
CH-1211 Genève 23
Switzerland
**www.cern.ch/it**

# Week-4: setup

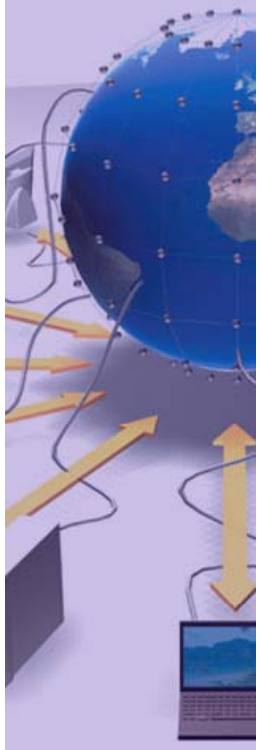





CERN IT Department
CH-1211 Genève 23
Switzerland
**www.cern.ch/it**

# Exercise





- T0 load generator
  - **Red**: runs at 100% of nominal rate
    - 14 hours of data taking at 200Hz, dispatched in 24h
    - Distributes data according to MoU (ADO everywhere ..)
  - **Blue:** runs at 100% of nominal rate
    - Produces only ESD and AOD
    - Distributed AOD and ESD proportionally to MoU shares
- T1s:
  - receive both **red** and **blue** from T0
  - Deliver **red** to T2s
  - Deliver **red** ESD to partner T1 and **red** AOD to all other T1s

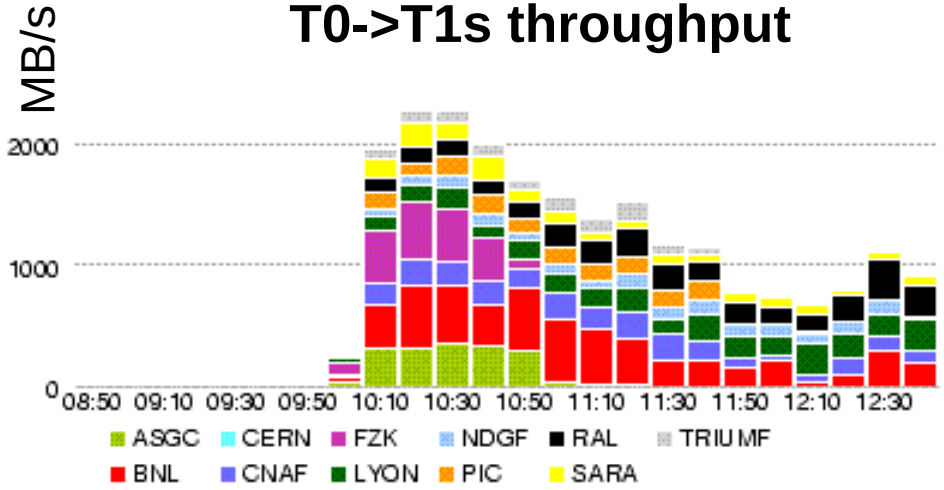
CERN IT Department
CH-1211 Genève 23
Switzerland
**www.cern.ch/it**

# Transfer ramp-up

CERN IT Department

Test of backlog recovery
First data generated over 12 hours and subscribed in bulk

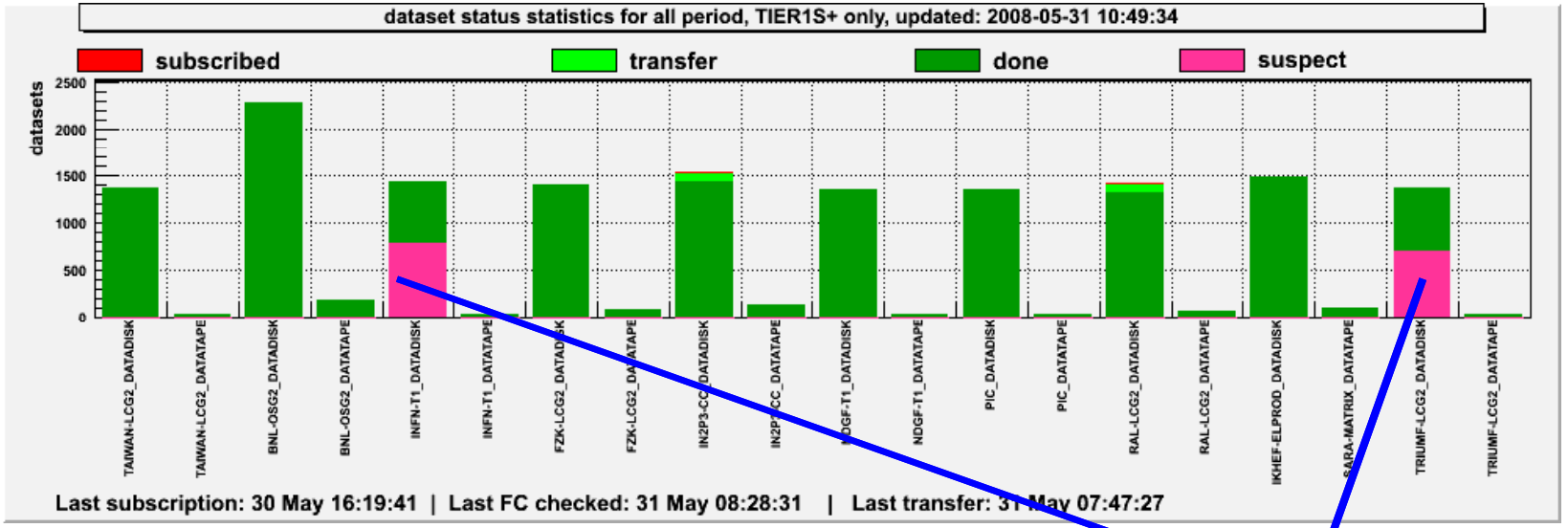
***12h backlog recovered in 90 minutes!***




| | Transfers | | | Registrations | | Errors | |
|---|---|---|---|---|---|---|---|
| Cloud | Efficiency | Throughput | Successes | Datasets | Files | Transfer | Registration |
| ASGC | 100% | 219 MB/s | 300 | 46 | 300 | 0 | 0 |
| BNL | 100% | 471 MB/s | 597 | 10 | 597 | 0 | 0 |
| CERN | 0% | 0 MB/s | 0 | 0 | 0 | 0 | 0 |
| CNAF | 100% | 195 MB/s | 196 | 17 | 196 | 0 | 0 |
| FZK | 100% | 229 MB/s | 331 | 40 | 329 | 0 | 0 |
| LYON | 99% | 147 MB/s | 155 | 9 | 156 | 2 | 0 |
| NDGF | 100% | 83 MB/s | 98 | 22 | 98 | 0 | 0 |
| PIC | 100% | 132 MB/s | 156 | 19 | 156 | 0 | 0 |
| RAL | 99% | 154 MB/s | 152 | 17 | 152 | 1 | 0 |
| SARA | 100% | 132 MB/s | 207 | 16 | 208 | 0 | 0 |
| TRIUMF | 100% | 105 MB/s | 94 | 26 | 92 | 0 | 0 |

CERN IT Department
CH-1211 Genève 23
Switzerland
**www.cern.ch/it**

# Week-4: T0->T1s data distribution




| Datasets | Total Files in datasets | Last Subscription | LFC Checked | Last Transfer |
|---|---|---|---|---|
| 691 | 9752 | May 30 17:06:39 | May 31 08:12:31 | May 31 08:12:30 |

| Tier1 | Datasets | Total Files in datasets | Total CpFiles in datasets | Completed | Transfer | Subscribed |
|---|---|---|---|---|---|---|
| BNL | 549 | 8170 | 8170 | 549 | 0 | 0 |
| FZK | 442 | 3400 | 3097 | 422 | 9 | 11 |
| IN2P3 | 432 | 3528 | 3432 | 426 | 0 | 6 |
| INFN | 464 | 3530 | 3530 | 464 | 0 | 0 |
| NDGF | 477 | 4033 | 4137 | 472 | 0 | 0 |
| PIC | 483 | 4046 | 4044 | 482 | 0 | 1 |
| RAL | 505 | 5013 | 4900 | 485 | 18 | 2 |
| SARA | 421 | 3137 | 3136 | 420 | 0 | 1 |
| TAIWAN | 470 | 4050 | 4036 | 464 | 5 | 1 |
| TRIUMF | 488 | 4221 | 4120 | 477 | 10 | 1 |

**Suspect Datasets**
***Datasets is complete (OK) but double registration***

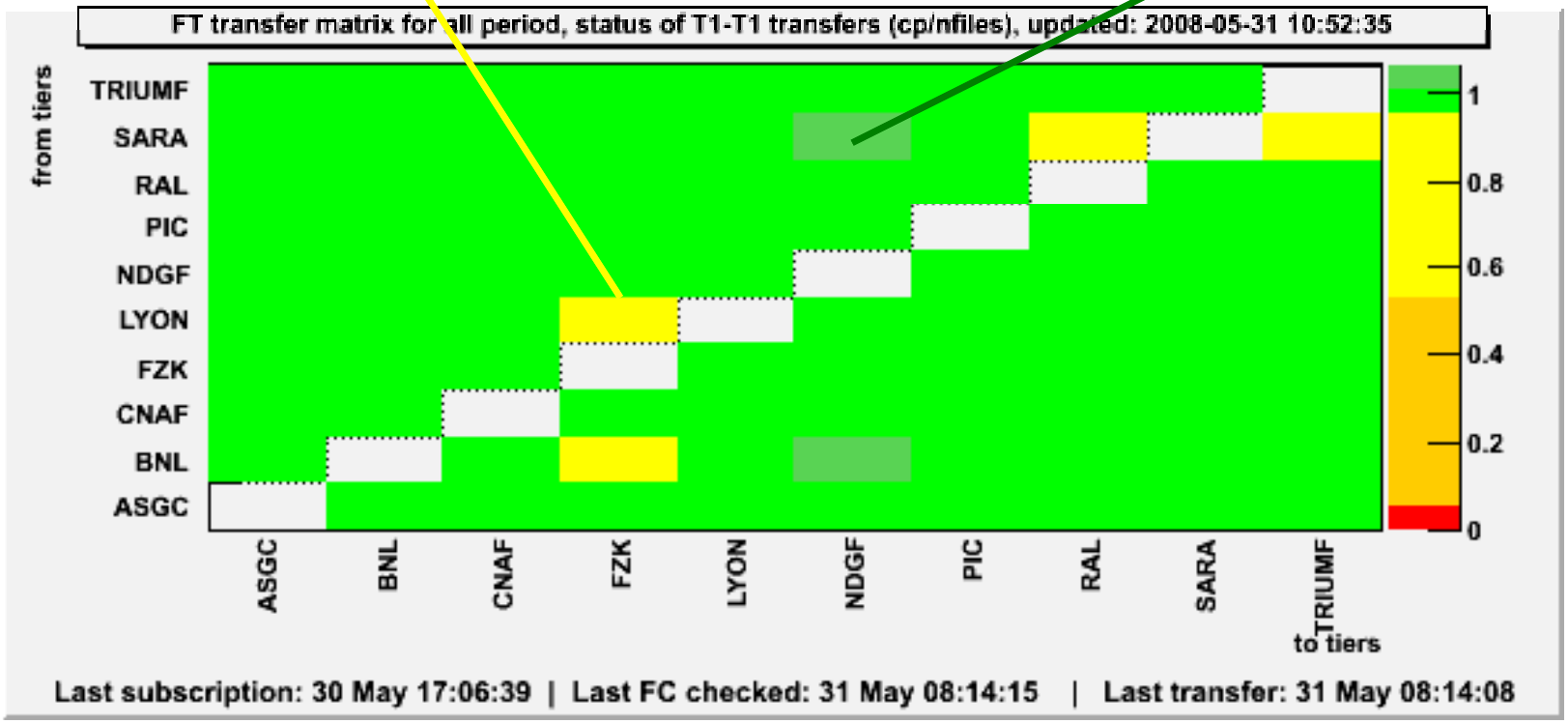
**Incomplete Datasets**
***Effect of the power-cut at CERN on Friday morning***


CERN IT Department
CH-1211 Genève 23
Switzerland
**www.cern.ch/it**

# Week-4: T1-T1 transfer matrix

YELLOW boxes
Effect of the power-cut

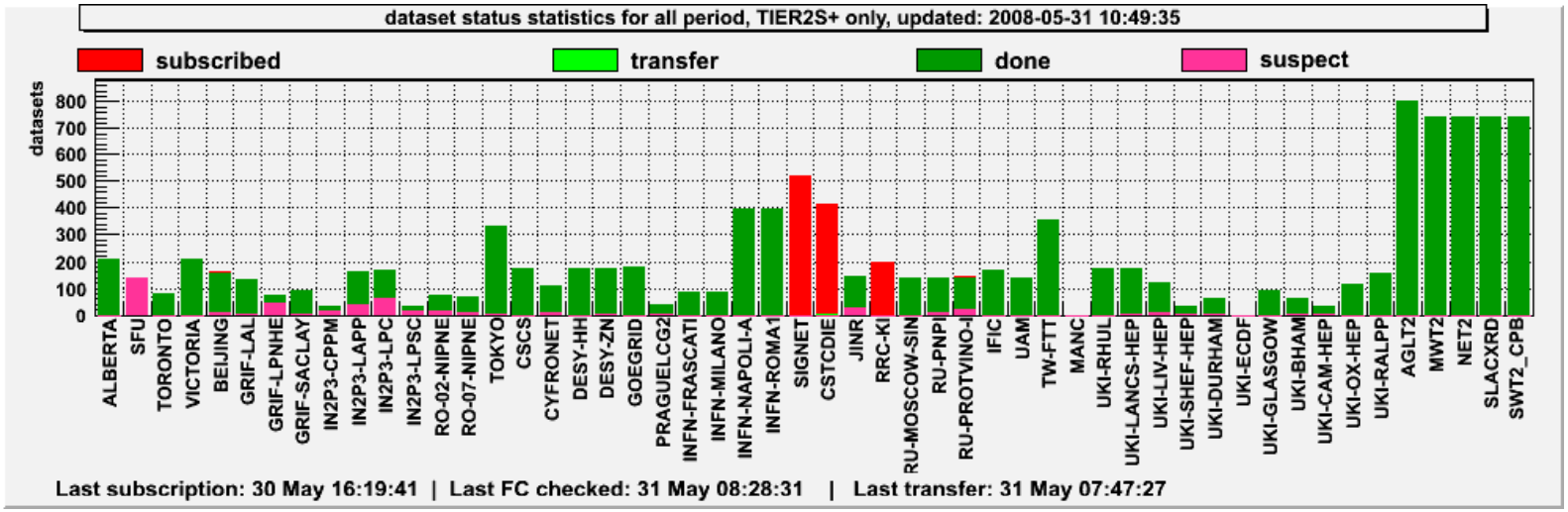
DARK GREEN boxes
Double Registration problem




**Compare with week-2 (3 problematic sites)**
**Very good improvement**

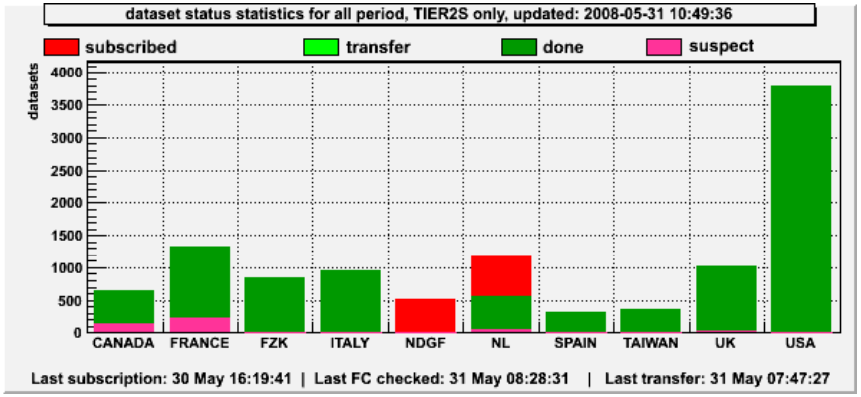
CERN IT Department
CH-1211 Genève 23
Switzerland
**www.cern.ch/it**

# Week-4: T1->T2s transfers




**SIGNET: ATLAS DDM configuration issue (LFC vs RLS)**

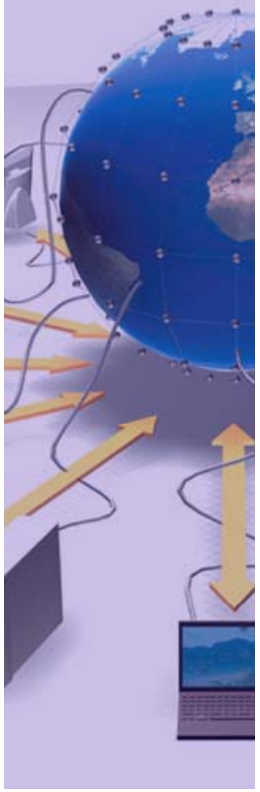
**CSTCDIE: joined very late. Prototype.**



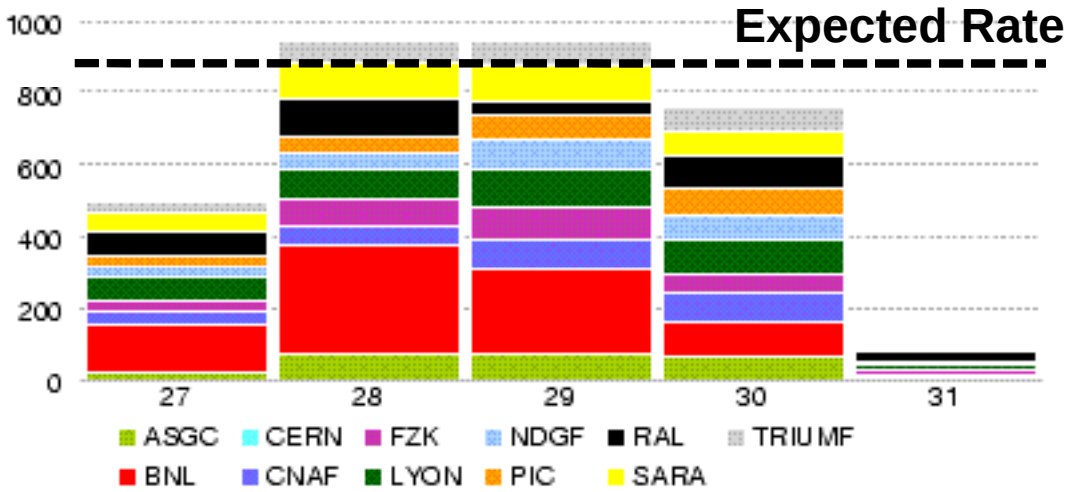

Many T2s oversubscribed
(should get 1/3 of AOD)

CERN IT Department
CH-1211 Genève 23
Switzerland
**www.cern.ch/it**

# Throughputs



MB/s




***T0->T1 transfers***

Problem at load generator on 27th

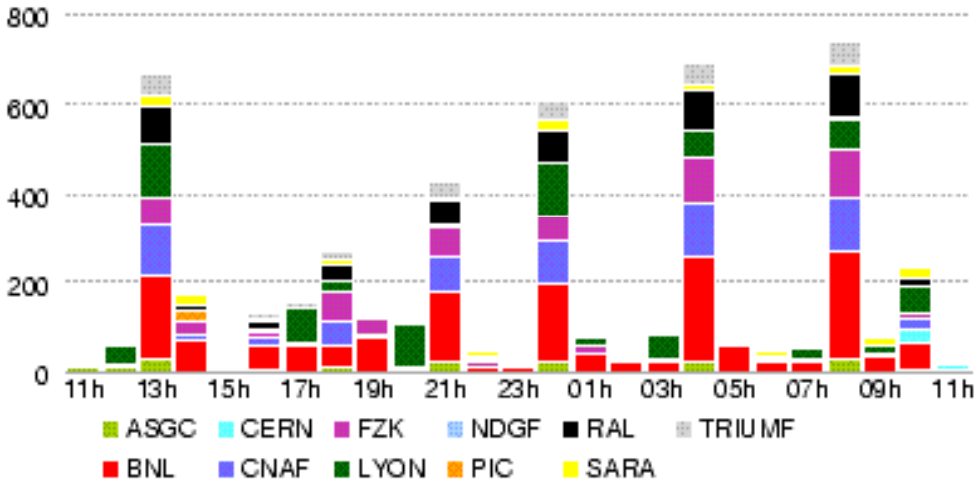
Power-cut on 30th

***T1->T2 transfers***

show a time structure

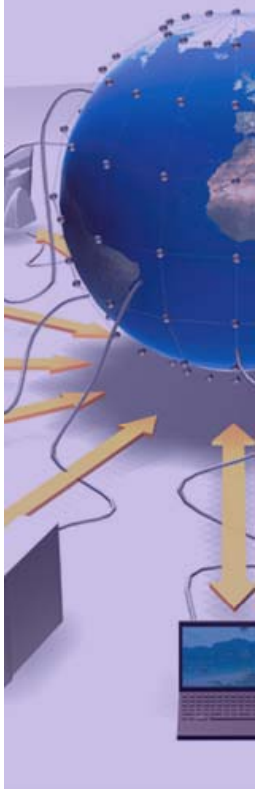
Datasets subscribed:
-upon completion at T1 -every 4 hours

MB/s

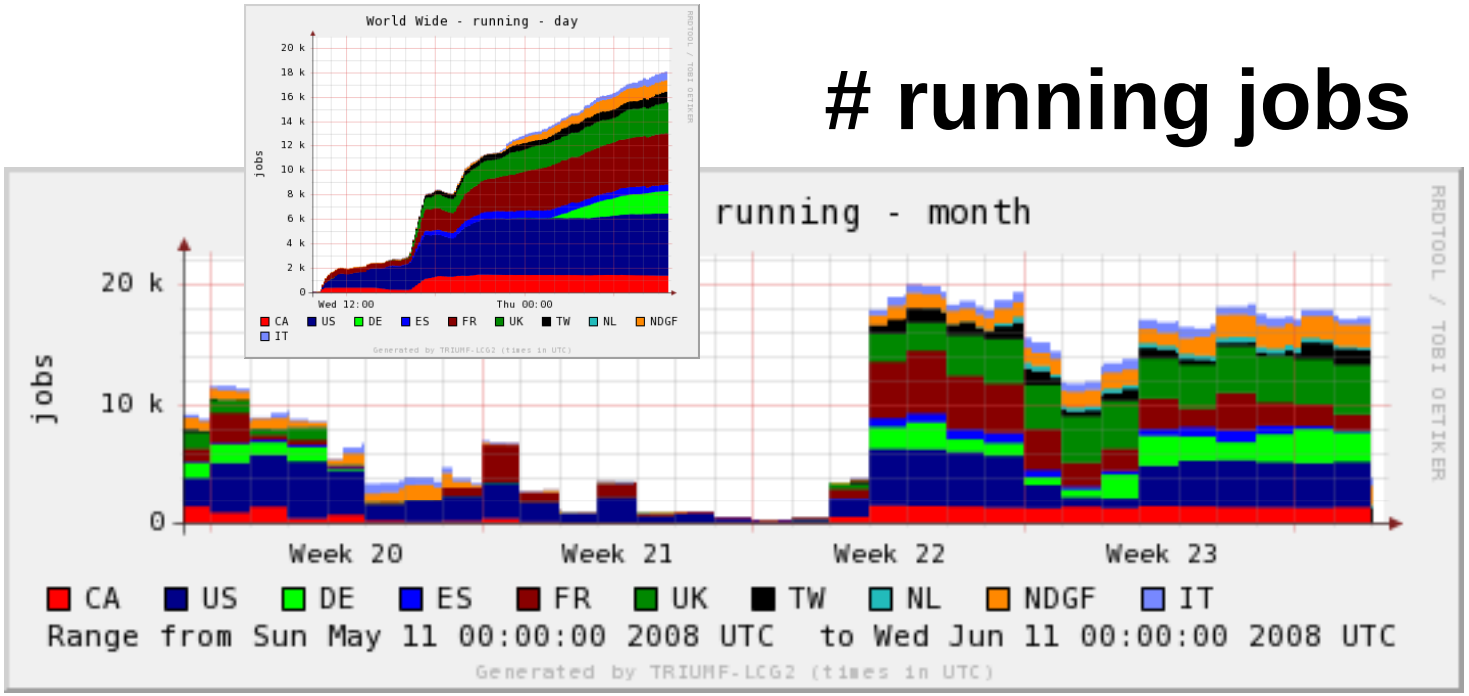



CERN IT Department
CH-1211 Genève 23
Switzerland
**www.cern.ch/it**

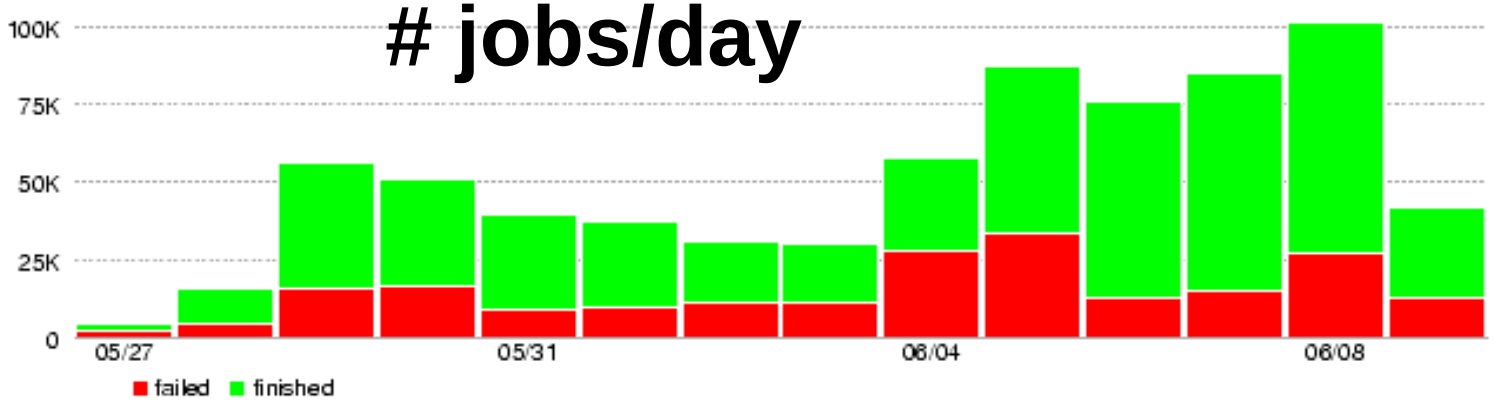
# Week-4 and beyond: Production



## # running jobs

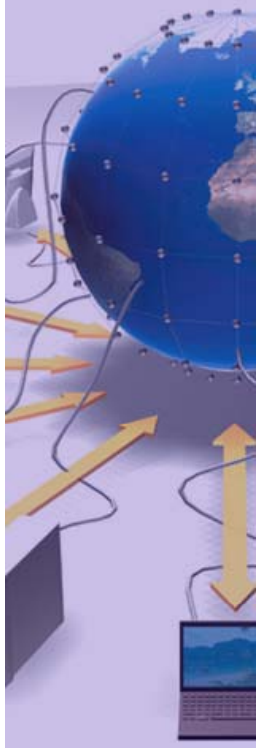



## # jobs/day




CERN IT Department
CH-1211 Genève 23
Switzerland
**www.cern.ch/it**

# Week-4: metrics



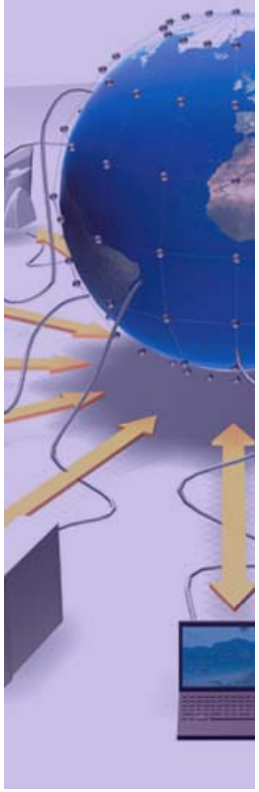
- We said:
  - T0->T1: sites should demonstrate to be capable to import 90% of the subscribed datasets (complete datasets) within 6 hours from the end of the exercise
  - T1->T2: a complete copy of the AODs at T1 should be replicated at among the T2s, withing 6 hours from the end of the exercise
  - T1-T1 functional challenge, sites should demonstrate to be capable to import 90% of the subscribed datasets (complete datasets) for within 6 hours from the end of the exercise
  - T1-T1 throughput challenge, sites should demonstrate to be capable to sustain the rate during nominal rate reprocessing i.e. F*200Hz, where F is the MoU share of the T1.
- Every site (cloud) met the metric!
  - Despite power-cut
  - Despite “double registration problem”
  - Despite competition of production activities

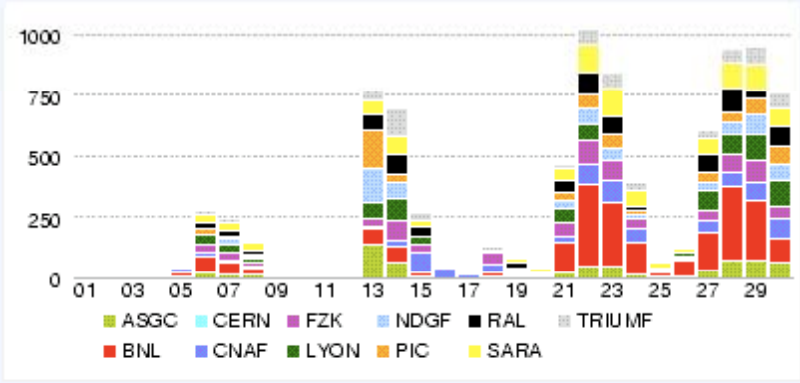
CERN IT Department
CH-1211 Genève 23
Switzerland
**www.cern.ch/it**

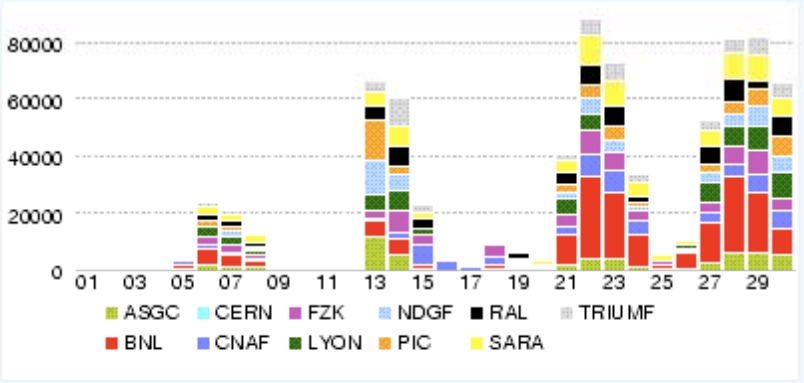
# All month activity

CERN IT Department

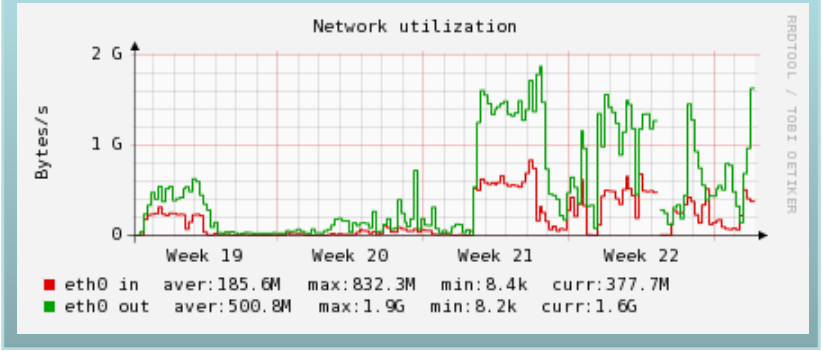
This includes both CCRC08 and detector commissioning

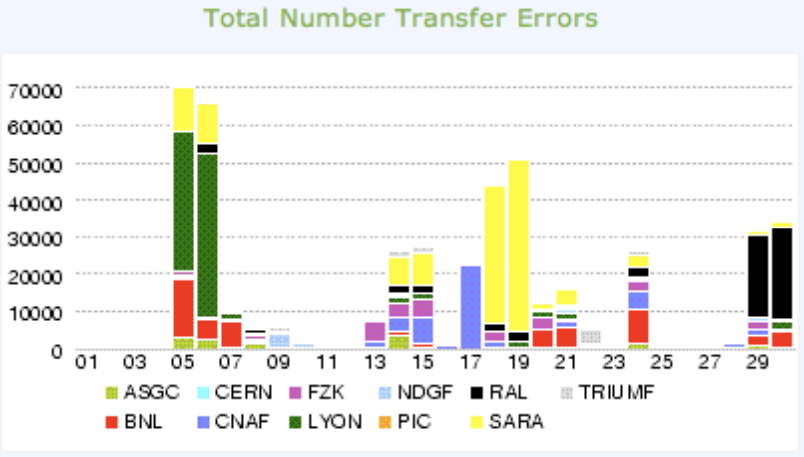




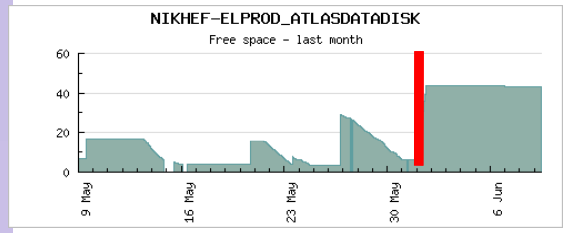

**CASTOR@CERN stress tested**

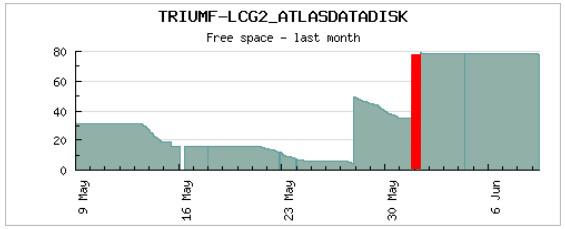


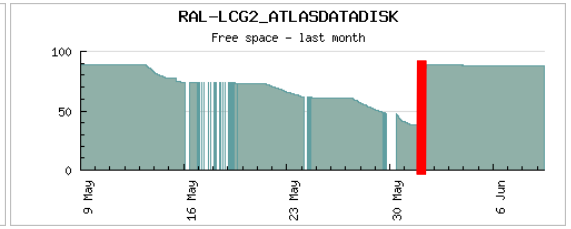



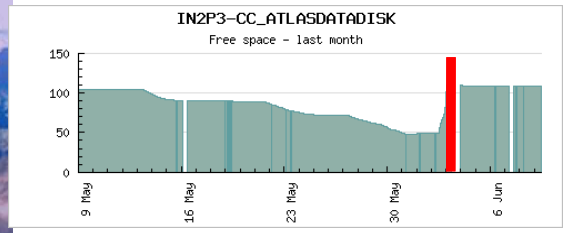

CERN IT Department
CH-1211 Genève 23
Switzerland
**www.cern.ch/it**

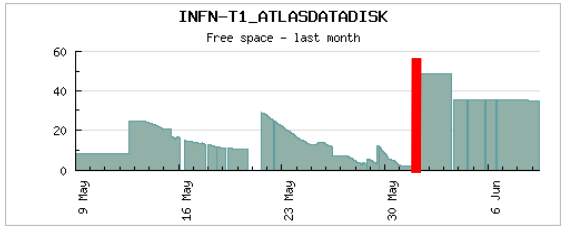
# Disk Space (month)

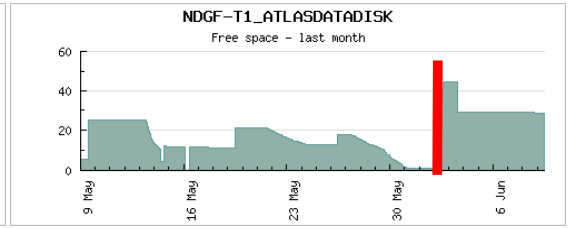


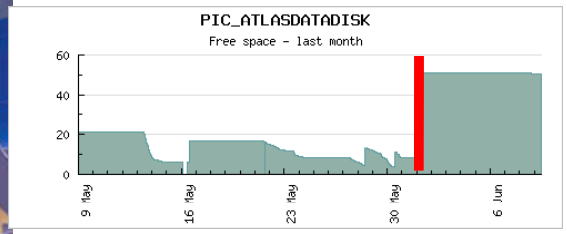



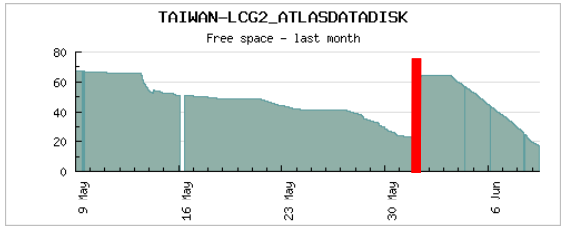



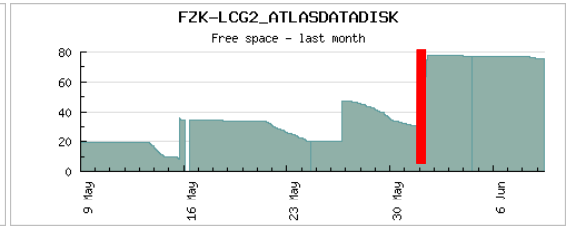



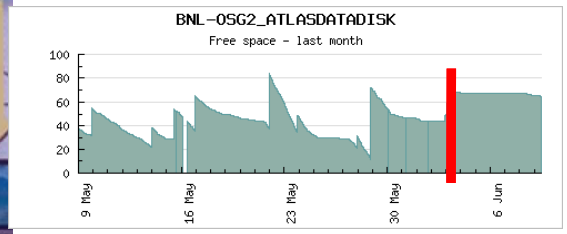




ATLAS “moved” **1.4PB** of data in May 2008

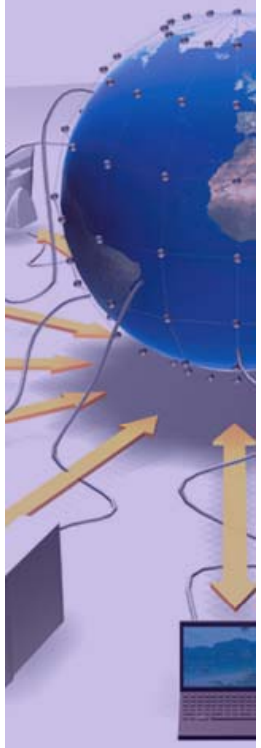
***1PB deleted in EGEE+NDGF in << 1day***

Order of **250TB deleted in OSG**

Deletion agent at work. Uses SRM+LFC bulk methods.
Deletion rate is more than good (but those were big files)

CERN IT Department
CH-1211 Genève 23
Switzerland
**www.cern.ch/it**

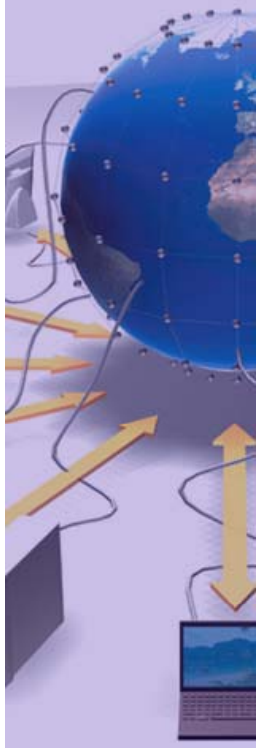
# Storage Issues: CASTOR



- "Too many threads busy with Castor at the moment"
  - SRM can not submit more requests to the CASTOR backend
    - In general can happen when CASTOR does not cope with request rate
    - Happened May 9th and 12th at CERN, sometimes at T1s
  - Fixed optimizing performance of stager_rm
    - "Hint" in Oracle query has been introduced
- Nameserver overload
  - Synchronization nameserver-diskpools at the same time of DB backup
  - Happened May 21st, fixed right after, did not occur again
- SRM fails to contact Oracle BE
  - Happened May 5th, 15th, 21st, 24th
  - Very exhaustive post mortem
    - https://twiki.cern.ch/twiki/bin/view/FIOgroup/PostMortemMay24
  - Two "fabric level" solutions have been implemented
    - Number of Oracle sessions on the shared database caped to avoid overload. SRM server and daemon thread pool sizes reduced to match max number of sessions
    - New DB hardware has been deployed
  - See talk from Giuseppe Lo Presti yesterday
  - Problem did not reappear after that.

CERN IT Department
CH-1211 Genève 23
Switzerland
**www.cern.ch/it**

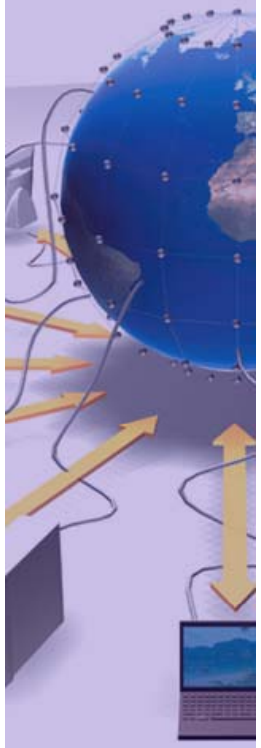
# Storage Issues: dCache



- Some cases of PNFS overload
  - FZK for the all T1-T1 test
  - Lyon and FZK during data deletion
  - BNL in Week-1 during data deletion (no SRM)
- Issues with orphan files in SARA not being cleaned
- Different  issues when disk pool is full/unavailable
  - Triumf in Week-2, PIC in Week-3
- The SARA upgrade to the latest release has been very problematic
  - General instability
  - PreStaging stopped working
    - dCache issue? GFAL issue? Whatever…
- More integration tests are needed, together with a different deployment strategy.

CERN IT Department
CH-1211 Genève 23
Switzerland
**www.cern.ch/it**

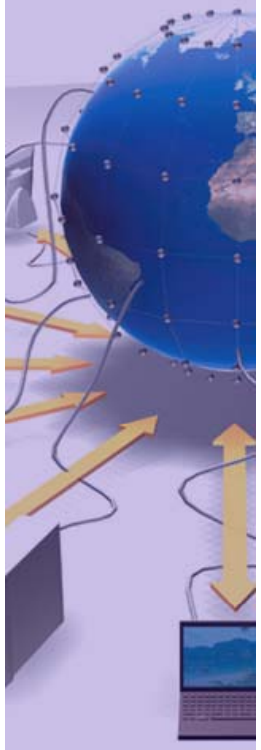
# Storage issues: StoRM and DPM



- StoRM
  - Problematic Interaction between gridftp (64 KB rw buffer) and GPFS (1 MB rw buffer)
    - Entire block re-written if #streams > #gridFTP servers
    - Need to limit FTS to 3 streams per transfer
  - Solutions:
    - Upgrade griFTP servers to SLC4
      - 256KB write buffer
      - More performing by factor 2
    - Deploy more (performing) hardware
    - Could push up #streams to 10
- DPM
  - No observed instability for Nikhef instance
  - Comments from T2s?

CERN IT Department
CH-1211 Genève 23
Switzerland
**www.cern.ch/it**

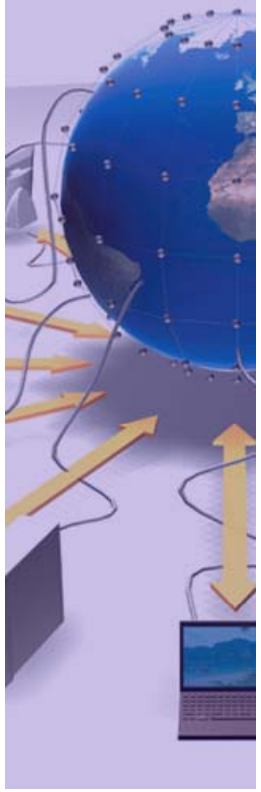
# More general issues

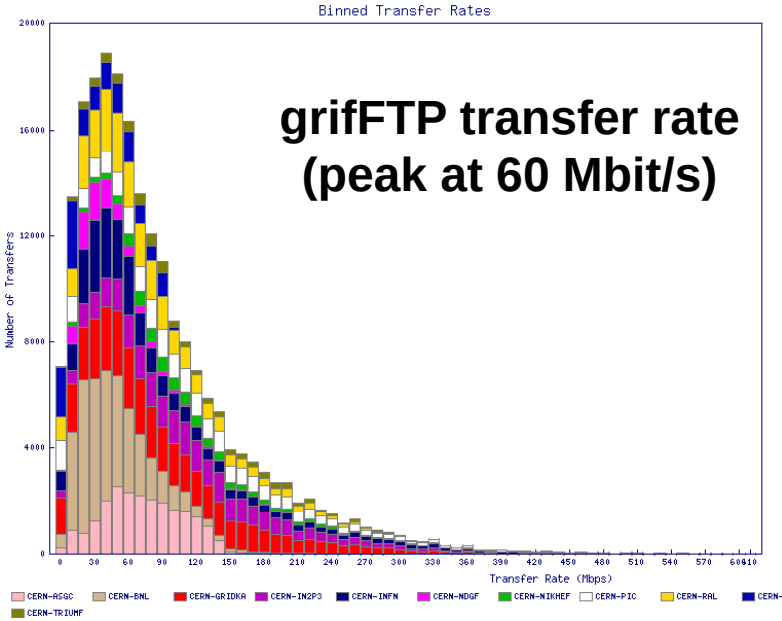


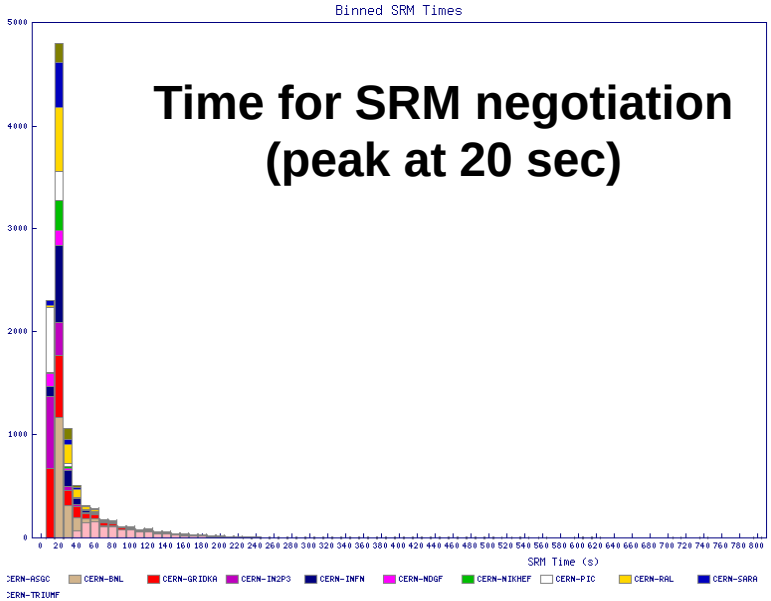
- **Network**
  - In at least 3 cases, a network problem or inefficiency has been discovered
    - BNLtoCERN, TRIUMFtoCERN, RALtoIN2P3
    - Usually more a degrade than failure … difficult to catch
  - How to prevent this?
    - Iperf server at CERN and T1s in the OPN?
- **Storage Elements**
  - In several cases the storage element “lost” the space token
    - Is this effect of some storage reconfiguration? Or can happen during normal operations?
    - In any case, sites should instrument some monitoring of space token existence
    - Hold on to your space tokens!
- **Power cut at CERN**
  - ATLAS did not observe dramatic delays in service recovery
  - Some issues related to hardware failures

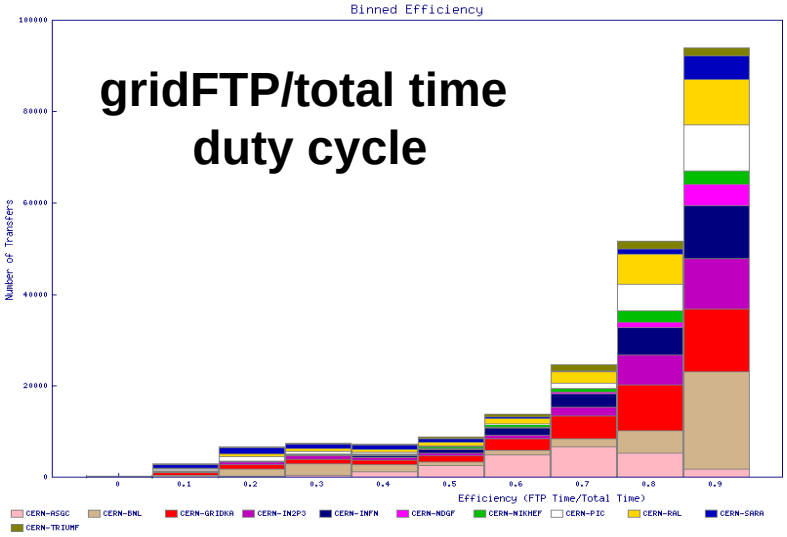
CERN IT Department
CH-1211 Genève 23
Switzerland
**www.cern.ch/it**

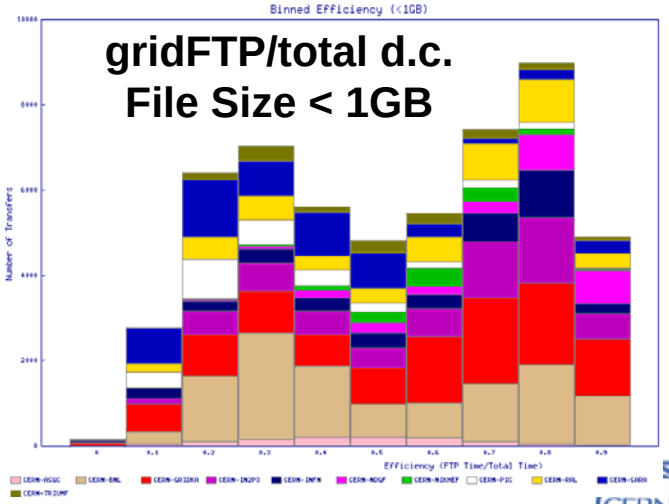
# Analysis of FTS logs in week 3 (successful transfers only)





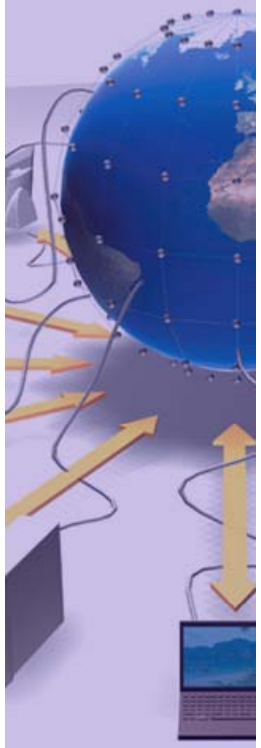




CERN IT Department
CH-1211 Genève 23
Switzerland
**www.cern.ch/it**

**From Dan Van Der Ster**

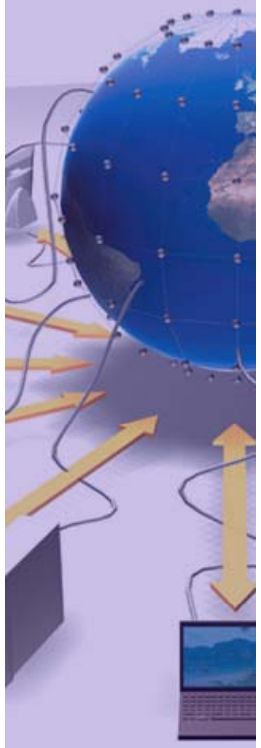
# More outcomes from CCRC08



- NDGF used **LFC instead of RLS**
  - No issues have been observed
  - Much simpler for ATLAS central operations
  - NDGF is migrating to LFC for production activities
    - Well advance migration plan
- CNAF tested a **different FTS configuration**
  - Two channels from T0 to CNAF
    - One for disk, one for tape, implemented using 2 FTS servers.
  - Makes sense if
    - DISK and TAPE endpoints are different at T1 or show very different performances
    - You assume SRM is the bottleneck and not the network
  - For CNAF, it made the difference
    - 90MB/s to disk + 90MB/s to tape in week 4
  - Where to go from here?
    - Dealing with 2 FTS servers is painful. Can we have 2 channels connecting 2 sites?
    - Probably needs non trivial FTS development

CERN IT Department
CH-1211 Genève 23
Switzerland
**www.cern.ch/it**

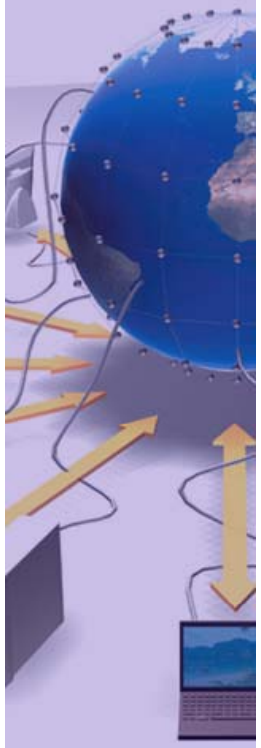
# Support and problem solving



- For CCRC08, ATLAS used **elog** as primary placeholder for problem tracking
  - There is also ATLAS elog for internal issues/actions
- Beside elogs, email is sent to **cloud mailing list** + atlas **contact at the cloud**
- In addition GGUS ticket is submitted
  - For traceability
- ATLAS follows a strict regulation for ticket severity
  - TOP PRIORITY: problem at T0, blocking all data export activity
  - VERY URGENT: problem at T1, blocking all data import
  - URGENT: degrade of service at T0 or T1
  - LESS URGENT: problem at T2 or observation of already solved problem at T0 or T1
- Shifters (following regular production activities) use **GGUS** as main ticketing system

CERN IT Department
CH-1211 Genève 23
Switzerland
**www.cern.ch/it**

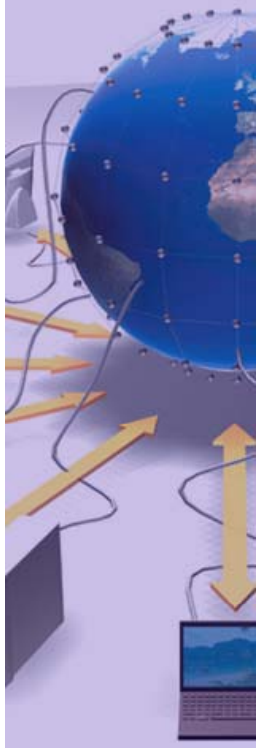
# Support and problem solving



- ATLAS submitted 44 elog tickets in CCRC08 (and possibly another 15/20 “private” request for help for small issues)
  - This is quite a lot … 2 problems per day
  - Problems mostly related to storage.
- I am impressed by the responsiveness of sites and service providers to VO requests or problems
  - Basically all tickets have been treated, followed, solved within 3 hours from problem notification
    - Very few exceptions
- The alarm mailing list (24/7 at CERN) has also been used on a weekend
  - From the ATLAS perspective it worked
  - But internally, the ticket followed an unexpected route
    - FIO followed up. We need to try again (may be we should not wait for a real emergency)

CERN IT Department
CH-1211 Genève 23
Switzerland
**www.cern.ch/it**

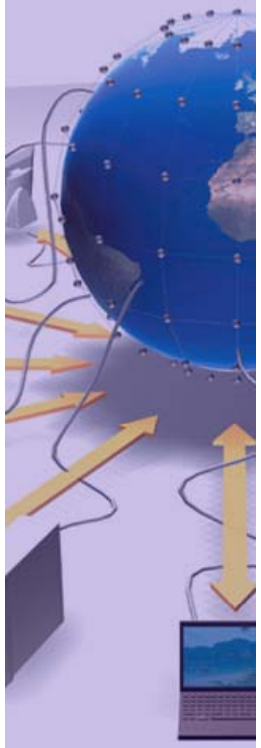
# ATLAS related issues



- The “double registration” problem is the main issue at the ATLAS level
  - Produces artificial throughput
  - Produces a disk space leak
  - Possibly caused by a variety of issues
    - But has to do with DDM-LFC interaction
      - http://indico.cern.ch/conferenceDisplay.py?confId=29458
  - Many attempts to solve/mitigate
    - Several version of ATLAS DDM Site Services deployed during CCRC08
    - Need to test current release
- The power cut has shown that
  - In ATLAS, several procedures are still missing
  - PITtoT0 data transfer “protocol” must be revisited
- Need to bring more people into the daily operation effort

CERN IT Department
CH-1211 Genève 23
Switzerland
**www.cern.ch/it**

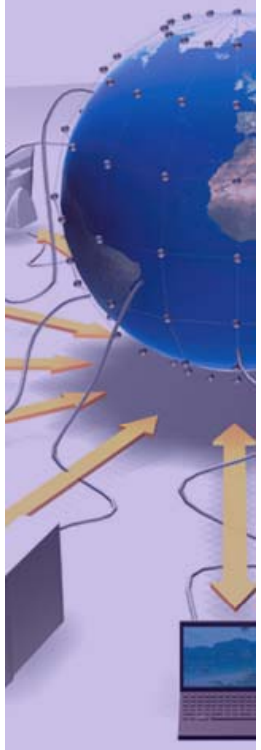
# What has not been tested in CCRC08



- **Reprocessing**
  - Tests are being carried along, but not in challenge mode
  - File staging from tape done via ATLAS pre-staging service
    - Using srm-bring-online via GFAL and srm-ls
  - Destination storage configuration has just been defined (T1D1 vs T1D0+pinning vs T1D0 with big buffers vs T1D0+T0D1)
- **Distributed Analysis**
  - Regular user analysis goes on every day
  - An “Analysis Challenge” has not been done
- Most likely, those will be the main test activities in the next months

CERN IT Department
CH-1211 Genève 23
Switzerland
**www.cern.ch/it**

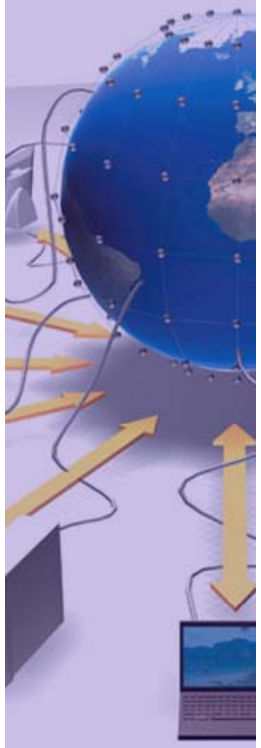
# Activity never stopped



- After CCRC08, activities did not stop
  - FDRII started the week after
- Few words about FDRII
  - Much less challenging than CCRC08 in terms of distributed computing
    - 6 hours of data per day to be distributed in 24h
  - Data distribution started at the end of the week
    - Three days of RAW data have been distributed in less than 4 hours
    - All datasets (RAW and derived) complete at every T1 and T2 (one exception for T2)
  - Unfortunately, a problem in the RAW file merging produced corrupted RAW files
    - Need to re-distribute the newly merged ones (and their derived)

CERN IT Department
CH-1211 Genève 23
Switzerland
**www.cern.ch/it**

# Conclusion



- **The data distribution scenario has been tested well beyond the use case for 2008 data taking**
- **The WLCG infrastructure met the experiment's requirements for the CCRC08 test cases**
- **Human attention will always be needed**

- Activity should not stop
  - ATLAS from now on will run continuous "heartbeat" transfer exercise to keep the system alive

CERN IT Department
CH-1211 Genève 23
Switzerland
**www.cern.ch/it**